What Constrains Spread Growth in Forecasts Initialized from Ensemble Kalman Filters?

Thomas M. Hamill and Jeffrey S. Whitaker

NOAA Earth System Research Lab, Physical Sciences Division

Boulder, Colorado

Corresponding author address:

Dr. Thomas M. Hamill
NOAA ESRL / PSD 1
325 Broadway, Boulder, CO 80305-3337
E-mail: tom.hamill@noaa.gov
Phone: (303) 497-3060
Fax: (303) 497-6449

ABSTRACT

The spread of an ensemble of weather predictions initialized from an ensemble Kalman filter may grow slowly relative to other methods for initializing ensemble predictions, degrading the skill of probabilistic predictions. Several possible causes of the slow spread growth were evaluated in perfect- and imperfect-model experiments with a 2-layer primitive equation spectral model of the atmosphere. The causes examined were covariance localization, the additive noise used to stabilize the model and parameterize the system error, and model error itself. For this experiment, the flow-independent additive noise was the biggest factor in constraining spread growth. Pre-evolving additive noise perturbations was tested as a way to make the additive noise more flow dependent. This improved the data assimilation and ensemble predictions both in the 2-layer model results and in a brief test of the assimilation of real observations into a global spectral primitive equation model.

1. Introduction

The ensemble Kalman filter, or "EnKF" (Evensen 2004; Houtekamer et al. 1998) and its variants (e.g., Hamill and Snyder 2000, Anderson 2001; Whitaker and Hamill 2002; Hunt et al. 2006) are being explored for their use in improving the accuracy of initial conditions and for initializing ensemble weather predictions. The EnKF produces an ensemble of parallel short-term forecasts and assimilations; background-error covariances from the ensemble are used in the data assimilation step. Introductory material on the EnKF is provided in Evensen (2006), Hamill (2006), and Ehrendorfer (2007). The technology behind the EnKF has matured to the point where it is used operationally for atmospheric data assimilation (Houtekamer et al. 2009) or is being tested actively with real data (e.g., Whitaker et al. 2004, 2008, 2009; Houtekamer et al. 2005; Houtekamer and Mitchell 2005; Compo et al. 2006; Miyoshi and Yamane 2007; Meng et al. 2008ab; Torn and Hakim 2008, 2009; Wang et al. 2008; Szunyogh et al. 2008; Zhang et al. 2009, Aksoy et al. 2009; Buehner et al. 2009ab).

The EnKF is now becoming a viable alternative to or complement of other advanced data assimilation schemes such as 4-dimensional variational data assimilation (4D-Var; Le Dimet and Talagrand 1986; Courtier et al. 1994; Rabier et al. 2000). A potential advantage that the EnKF may have for ensemble prediction is that an ensemble of initial conditions is automatically generated that, theoretically at least, have the proper characteristics for initializing ensemble forecasts (Kalnay et al. 2006). In comparison, an additional step is needed to create the ensemble of initial conditions when using 4D-Var

for the data assimilation. Hybridizations of these two methods are possible (Buehner et al. 2009ab).

Only a modest amount of experimentation has been performed on the characteristics of ensemble predictions initialized from EnKFs with real observations. Houtekamer et al. (2005) showed that in an earlier implementation of their EnKF, the spread of ensemble forecast perturbations was significantly smaller than the ensemble-mean forecast error (ideally, the two should be comparable in magnitude). Spread actually decreased during the first 12-24 h of the forecast. They attributed this in part to the structured noise added to each member (each a random draw consistent with the 3D-Var background-error statistics) used to address "system error" ensure that their EnKF maintained a consistency between ensemble spread and innovation statistics. They suggested also that their use of an overly diffusive forecast model, especially near the model top, unrealistically constrained spread growth.   More recently, Charron et al. (2009) reported greater spread growth when the use of the excessively diffusive model was eliminated.

Previously, Mitchell et al. (2002) had also demonstrated that the "covariance localization" applied in the EnKF to mute spurious long-distance covariances in the ensemble estimates (Houtekamer and Mitchell 2001; Hamill et al. 2001) introduced imbalances into the ensemble of analyzed states, which may also constrain error growth. Subsequently, Lorenc (2003), Buehner and Charron (2007), Bishop and Hodyss (2009ab) and Kepert (2009) have also discussed this effect and have suggested possible algorithmic modifications to remedy this.

What are the predominant mechanisms that constrain spread growth in ensemble filters? In addition to covariance localization and the additive noise, the forecast model may have a very different chaotic attractor (Lorenz 1993) than that of the natural atmosphere, known more simply as "model error." The data assimilation and short-range forecasts may produce a rapid oscillation of the model state back and forth, toward the observations and the atmosphere's attractor during the update step and back toward the model attractor during the forecast step. It is possible that this results in less projection of the perturbations onto the model's unstable manifold, and hence constrains perturbation growth. Other possibilities for the slow spread growth include the nature of effective data assimilation, which adjusts the background more toward the observations in the directions in phase space where background errors are large (and presumably spread growth is large). The analysis process naturally whitens the analysis-error spectrum relative to the background error spectrum (Hamill et al. 2002), decreasing the projection onto the growing modes. Another possible inhibitor to spread growth may be the incorrect specification of the observation-error covariances. A common assumption in data assimilation methods is that observations have uncorrelated errors, i.e., a diagonal observation-error covariance matrix. If in fact the observations have strong correlations but are assumed to have none, the resulting analysis may have a grid points with covariances that are too small in magnitude (Fig. 1). Consequently, ensembles of analyses contain an abundance small-scale, dynamically irrelevant noise.

In this manuscript we seek to understand some of the mechanisms for sub-optimal spread growth in ensemble Kalman filters. In particular, we will examine the effects of

covariance localization, additive error noise, and model error. We perform a simulation experiment with a simple, two-level primitive equation model, a model that hopefully is a realistic enough analog to shed light on approaches to be tried in modern-day numerical weather prediction models, but simple enough to permit the generation of a very large number of tests and many cases. The rest of the manuscript is organized as follows. Section 2 provides a brief review of the model and the data assimilation system used. Section 3 shows the results of simulation experiments under perfect-model conditions, and section 4 with an imperfect model. Section 5 provides results from experiments with a global numerical weather prediction model and real observations, and section 6 conclusions.

2. **The forecast model, data assimilation system, and experimental design**.

a. Forecast model.

The forecast model used in these experiments was virtually identical to the two-level spectral model of Lee and Held (1993), and a version of it with hemispheric symmetry was used for the ensemble data assimilation experiments in Whitaker and Hamill (2002). No hemispheric symmetry was imposed for these experiments. Here, the data assimilation experiments were run at T31 horizontal resolution, though for imperfect-model data assimilation experiments the nature run was computed at T42 resolution. The prognostic variables of the forecast model are baroclinic and barotropic vorticity, baroclinic divergence, and interface barotropic potential temperature. Barotropic divergence was set to zero, and baroclinic potential temperature was set to 10K. Lower-level winds were mechanically damped with an e-folding timescale of 4

days (4.5 days for T42 nature run). The baroclinic potential temperature was relaxed

back to a radiative equilibrium state with a pole-to-equator temperature difference of 80K

(74K for T42 nature run) with a timescale of 20 days. The radiative equilibrium profile

of Lee and Held (1993; eq. 3) was used. ∇^8 diffusion was applied to all the prognostic

variables, with the smallest resolvable scale damped with an e-folding timescale of 3 h.

Time integration proceeded with a 4th – order Runge-Kutta scheme with 18 time steps per

day (64 for the T42 nature run). The error doubling time of the T31 model was

approximately 2.4 days.

This model is obviously much simpler than the operational numerical weather

prediction models currently in use; the resolution is lower, there is no terrain, no land and

water, no atmospheric moisture. In fact, while this model is capable of supporting

internal gravity waves, it does not produce an external mode. These simplifications

should be kept in mind while interpreting the results and their implications for

operational numerical weather prediction.

b. Data assimilation methodology.

The ensemble square-root filter (EnSRF) of Whitaker and Hamill (2002) was used

for the data assimilation. The EnSRF, like other EnKF algorithms, consists of two steps

that are repeated, a set of short-range ensemble forecasts, and a data assimilation that uses

the short-range forecasts to estimate the background-error covariances for the ensemble

update. Assume that an ensemble of forecasts estimating the state at a particular time

when new observations are ready to be assimilated. The EnSRF algorithm separates the

EnKF update into an update of the mean and an update of the perturbations around the

mean. Observations are assimilated serially, so that the updated mean and perturbations

after the assimilation of the first observation are used as the background for the second

observation, and so on. Let $\overline{\mathbf{x}}^a$ denote the mean analysis state at the current time,

$\overline{\mathbf{x}}^b$ denote the mean background state, $\mathbf{y}^o$ denote the current observation, and $\mathbf{H}$ denote the

observation operator that converts the background state to the observation location and

type; here this operator is linear. The EnSRF update equations applied to this simplified

model and simplified observations are:

$$\overline{\mathbf{x}}^a = \overline{\mathbf{x}}^b + \mathbf{K}\left(\mathbf{y}^o - \mathbf{H}\overline{\mathbf{x}}^b\right) \qquad , \qquad (1)$$

where the Kalman gain $\mathbf{K}$ is

$$\mathbf{K} = \mathbf{P}^b\mathbf{H}^T\left(\mathbf{H}\mathbf{P}^b\mathbf{H}^T + \mathbf{R}\right)^{-1}. \qquad (2)$$

Here $\mathbf{R}$ denotes the observation-error variance and $\mathbf{P}^b$ the estimate of the background-

error covariance from the ensemble. This covariance matrix was not explicitly

calculated, but instead $\mathbf{P}^b\mathbf{H}^T$ was calculated in the EnSRF as a product

$$\mathbf{P}^b\mathbf{H}^T = \frac{1}{n-1}\sum_{i=1}^{n}\left(\rho \circ \mathbf{x}_i^{b'}\right)\left(\mathbf{H}\mathbf{x}_i^{b'}\right)^T \qquad , \qquad (3)$$

where $\mathbf{x}_i^{b'}$ is the ith of n member's deviation from the ensemble mean and ρ denotes the

Gaspari and Cohn (1999) correlation function. For each element of the state vector, the

distance between this element of the state vector and the current observation location is

computed. ρ then is a correlation vector; the correlations are approximately Gaussian-

shaped function, 1.0 at the observation location and tapering to zero at and beyond a user-

specified distance. Thus, $\rho \circ \mathbf{x}_i^{b'}$ computes the element-wise product of this correlation vector with the ith ensemble background state's perturbation. Similarly, $\mathbf{HP}^b\mathbf{H}^T$ is constructed without every explicitly computing $\mathbf{P}^b$:

$$\mathbf{HP}^b\mathbf{H}^T = \frac{1}{n-1}\sum_{i=1}^{n}\left(\mathbf{Hx}_i^{b'}\right)\left(\mathbf{Hx}_i^{b'}\right)^T \qquad . \qquad (4)$$

Equations (1) – (4) indicate how this implementation of the EnSRF updates the mean state to a new observation. Perturbations around the mean used a slightly different update, following Whitaker and Hamill (2002). Let $\mathbf{x}_i^{a'}$ denote the updated analysis perturbation for the ith member around the analyzed mean state. Then the update of the perturbations proceeded according to

$$\mathbf{x}_i^{a'} = \mathbf{x}_i^{b'} - \tilde{\mathbf{K}}\mathbf{Hx}_i^{b'} \qquad , \qquad (5)$$

where $\tilde{\mathbf{K}}$, the "reduced" Kalman gain, was calculated according to

$$\tilde{\mathbf{K}} = \left(1 + \sqrt{\frac{\mathbf{R}}{\mathbf{HP}^b\mathbf{H}^T + \mathbf{R}}}\right)^{-1} \qquad . \qquad (6)$$

By adding $\mathbf{x}_i^{a'}$ to $\bar{\mathbf{x}}^a$ for each member, an ensemble of analyzed states are reconstructed, and the full nonlinear forecast model is used to integrate each member forward to the next time when observations are available. This process is then repeated for the duration when observations are assimilated. If desired, at any time the ensemble of analysis states can be integrated forward for a longer period of time to produce an ensemble of weather forecasts.

In ensemble filters, covariances estimates from the ensemble are typically modified to stabilize the system and account for errors due to system errors such as model error or sampling (Hamill 2006, section 6.4.4). The covariance localization applied in eq. (3) is one form of stabilization. Two others were tested here. The first was "covariance inflation" (Anderson and Anderson 1999), whereby before the assimilation of the first observation, the deviation of every member of the ensemble around its mean was inflated by a constant:

$$\mathbf{x}_i^{b'} \leftarrow (1+\alpha)\mathbf{x}_i^{b'} \qquad , \tag{7}$$

where $\leftarrow$ denotes a replacement of the previous state perturbation on the right-hand side.

Another method of stabilization is "additive noise." Either before or after the update, structured noise is added to each ensemble member (e.g., Mitchell and Houtekamer 2000; Houtekamer et al. 2005, Hamill and Whitaker 2005). In this experiment, for the ith member, additive noise ε_i was added to the ith analyzed ensemble member $\mathbf{x}_i^a$ before integration forward to the next time when observations were available:

$$\mathbf{x}_i^a \leftarrow \mathbf{x}_i^a + \varepsilon_i \quad . \tag{8}$$

Presumably, this additive noise is a sample from the model-error covariance matrix $\mathbf{Q}$. How samples of additive noise were generated will be explained in the following section.

c. Experimental design.

Two sets of experiments were conducted, perfect-model and imperfect-model experiments. In each experiment the ensemble-mean error, ensemble spread, and ensemble spread growth was examined for parallel cycles testing a variety of stabilization techniques (localization, additive noise, covariance inflation). Unless mentioned otherwise, the ensemble size was n=50, and the same forecast model dynamics was used for each member; the model incorporated no stochastic physics, nor did it use multiple models.

In both the perfect and imperfect-model experiments, an observation network with 490 nearly equally spaced observation locations was used. The observations were located at the nodes of a geodesic grid. At each location, observations were created for the barotropic potential temperature and the u- and v- wind components at the two model levels, 250 and 750 hPa. Observations were created by interpolating the true state to the observation location and adding random, independent, normally distributed observation errors. Errors had zero mean and variances of 1 K^2 and 1.0 m^2s^{-2} for potential temperature and winds, respectively. The nature run for generating the true state was produced by starting the forecast model from a random perturbation superimposed on a resting state, integrating for 350 days and discarding the first 200 days. Observational data was assimilated over the 150 days, with an update to new observations every 12 h. In the computation of assimilation statistics, the first 25 days of data assimilation were discarded due to transient effects.

Errors, spread, and spread growth were measured in the total-energy norm:

$$\| \cdot \| = \frac{\sqrt{\frac{1}{2} \int\limits_A \left[u^2 + v^2 + \frac{c_p}{T_{ref}} T^2 \right] dA}}{\int\limits_A dA} \tag{9}$$

where c_p is the specific heat capacity of dry air at constant pressure (1004 J kg^{-1} K^{-1}), T_{ref} = 300 K, and the integrals were performed over the earth's surface area. Statistics were calculated separately for each data assimilation cycle and then averaged over all of the 125 remaining days.

Different types of additive noise were used for the perfect- and imperfect-model simulation experiments. In the perfect-model experiments, the additive noise was generated as follows. Using the nature run, a time series of differences between the model state at time t and time $t + 24$ h were calculated. At any particular assimilation time, 50 random samples of these differences were chosen from the time series, without replacement. The mean state of the 50 samples was computed and subtracted from the 50 random samples. Denote the ith sample as ζ_i. These additive noise samples were then scaled by a constant α and added to the ensemble of *analysis states*, i.e., in eq. (8) $\varepsilon_i = \alpha \zeta_i$. In some experiments, α was pre-specified. In other "adaptive additive" experiments, α was computed in order to provide the correct innovation statistics. It can be shown readily that the expected value $\langle \cdot \rangle$ of the difference between the observations and the ensemble-mean analysis ought to be matched by the sum of the observation-error variance and the analysis perturbation variance at the observation locations:

$$\left\langle \left(\mathbf{y} - \mathbf{H}\overline{\mathbf{x}}^a \right)^2 \right\rangle = \left\langle \mathbf{R}^2 \right\rangle + \left\langle \left(\mathbf{H}\mathbf{x}_i^{a'} \right)^2 \right\rangle \tag{10}$$

Hence, the magnitude of α was, in most perfect-model experiments, chosen so that

$$\left\langle \left(\mathbf{y} - \mathbf{H}\overline{\mathbf{x}}^{a}\right)^{2}\right\rangle = \left\langle \mathbf{R}^{2}\right\rangle + \left\langle \left(\mathbf{H}\left(\mathbf{x}_{i}^{a'} + \alpha\zeta_{i}\right)\right)^{2}\right\rangle \qquad . \qquad (11)$$

When the right-hand side of eq. (10) was larger than the left before the addition of any noise, α was set to 0.0.

The perfect-model additive noise typically had large amplitudes in the mid-latitude storm tracks and much smaller amplitudes elsewhere (not shown). Since the imperfect-model nature run used a higher-resolution model (T42), a different pole-to-equator temperature difference (74K), and a different mechanical damping time scale (4.5 days), one would expect systematic differences between the forecast model's and the nature run's climatologies beyond the storm track. Hence, imperfect-model additive noise was designed to sample these possible differences in forecast and true model states. To produce such additive noise samples, multiple T31 nature runs were created, each using a different pole-to-equator temperature difference and damping time scale. Pole-to-equator temperature differences ranged from 74 to 83 K, and damping timescales ranged from 3 to 5 days. Figure 2 shows the zonal-mean profiles of the upper- and lower-level u-wind component, and interface potential temperature for the forecast model nature run, the T42 nature run, and the set of perturbed T31 nature runs. 400 random model states were extracted from the set of perturbed T31 nature runs. The 50 samples of additive noise at any particular update time during a data assimilation experiment were drawn randomly from the 400 perturbed states, without replacement. The mean state of these 50 members was then calculated and subtracted from each to create 50 perturbed

states. These samples of additive noise were typically reduced in amplitude, the magnitude of α specified in the experiment.

3. **Perfect-model experiment results**.

Figure 3 shows error and spread for perfect-model experiments. Here, multiple parallel cycles of the EnSRF were conducted, varying the covariance localization across a range of length scales and stabilizing the data assimilation either with 2% covariance inflation or additive noise. The magnitude of the additive noise was determined adaptively each update step using the procedure described in section 2c. Several characteristics of the spread and error are notable. First, errors were strongly affected by the covariance localization length scale. Very small length scales produced analyses with larger errors, and analysis-error minima were found with localization length scales of approximately 10,000 km for the additive-noise and 15,000 km for the covariance inflation. Similar effects of localization on error were previously demonstrated in Houtekamer and Mitchell (1998), Houtekamer and Mitchell (2001), and Hamill et al. (2001).

Overall, the covariance inflation simulations had much less error and a greater consistency between spread and error than the adaptive additive error simulations. This raises two questions: first, why do the adaptive additive error simulations have more spread than error? And why did they have larger analysis errors? As for why the there was an inconsistency in spread, a likely reason for this was that the amount of adaptive additive error was chosen to ensure a consistency between spread and error *at the observation locations*. However, the spread and error shown in Fig. 3 were calculated

globally, at points both near to and relatively far from the observations. As the localization length scale was shortened, the potential corrective effect of an observation was limited more and more to nearby the observation, while away from it the analysis continued to reflect the influence of the prior and preserved the prior's spread.

Why did the adaptive additive error simulations have more error? Figure 4 shows the growth rate of spread during the 12-h between updates to the observations. Spread generally grew more slowly in the adaptive additive error simulations; the adaptive noise was not dynamically conditioned to the flow of the day, while covariance inflation preserved the flow-dependent structures. The exceptions to the higher errors with additive noise were at the small localization length scales. For these parameter values, the EnSRF was stabilized purely by the covariance localization, and the adaptive additive noise typically consisted of no noise at all. Consequently, the growth rates were more similar to those from the covariance-inflation simulations.

Other characteristics of spread growth are also evident in Fig. 4. At large localization length scales, the rate of growth of spread for the covariance inflation simulations was approximately equal to that from a 400-member simulation that was stabilized by 1% inflation but which utilized no localization whatsoever. Based on this, the assumption is that the growth rate of ~1.20 was taken to represent an approximate upper-limit of the possible spread growth rate in this model with an EnSRF. The use of a short localization radius did decrease the rate of growth of spread somewhat, approximately 2 to 3 percent for the covariance inflation simulations relative to the large localization radii. Spread growth was smaller for the adaptive additive error simulations,

15

but spread growth did not increase as strongly with increasing localization radius as did the covariance inflation. This was probably because the greater growth rate from less localization was somewhat counteracted by the slower growth from the application of greater amounts of additive noise to stabilize the filter.

Overall, in this model, covariance localization had only a small effect on the rate of growth, decreasing growth by a few percent. The imbalances introduced from the additive noise approximately doubled that effect. The magnitudes of these changes in growth rate may have been an artifact of this simple model and may not be realistic of what may occur in real numerical weather prediction models. Such models may support additional unbalanced modes (e.g., external gravity waves) and may both generate noise and organize it much more readily due to the presence of moist convection (Zhang et al. 2003).

4. **Imperfect-model results**.

Figure 5 provides error, spread, and spread growth statistics for experiments with a variety of combinations of covariance inflation magnitude and covariance localization length scale. Errors in general were much higher than the perfect-model results, and much more inflation and much tighter localization were needed to stabilize the filter. The minimum error was obtained at a localization length scale of 3000 km and an inflation rate of 50 percent. However, the analysis spread was smaller than the error for this length scale/inflation. Errors increased dramatically as the covariance inflation amount was lowered, while spread decreased and spread growth increased; this was a sign that filter divergence was occurring. There were some combinations of large inflation and tight

localization where the filter was numerically unstable; when a tight localization was used, model-state perturbations away from the observations occasionally experienced an uncontrolled growth of spread that led to numerical instabilities.

Aside from when very small inflation was applied, spread growth was decreased. In fact, spread decayed on average during the 12 h for the length scale and inflation that produced the minimum of error. Figure 6 illustrates the challenges of tuning the covariance inflation to produce the optimal spread. Here, zonal and time-averaged spread and error are plotted for the minimum-error inflation rate/length scale. Spread was generally smaller than error, but was greater than error for tropical temperatures. When spread was further increased, temperature and low-level wind errors increased in the tropics (not shown), indicating that the drastic inflation was degrading the correlation structures in the ensemble.

Unlike the perfect-model experiments, the imperfect-model experiments produced slightly smaller analysis errors when the EnSRF was stabilized with additive noise instead of covariance inflation (Fig. 7). When more additive noise was applied to stabilize the filter, the analysis spread increased, as expected. Unexpectedly, the spread growth rate varied only slightly no matter how much additive noise was applied. Relative to the perfect-model experiments where spread growth was as large as 1.16, the spread growth for these imperfect model experiments was lower, with a maximum of approximately 1.12. Figure 8 shows that there was a greater consistency between spread and error across latitudes and variables, though temperature spread in the mid-latitude storm track was too large by nearly an order of two.

With the results presented thus far, it is difficult to determine whether the decrease in spread growth relative to perfect-model experiments can be attributed primarily to the additive noise or to the effects of model error. To better quantify the potential effects of additive noise, an additive noise perturbation was added to a nature run from the T31 forecast model, and size of the perturbation was calculated in the energy norm as the control and perturbed forecasts were integrated to 4 days lead. This process was repeated over 11 different case days, equally spaced every 12.5 days during the nature run. The zonal and sample-average growth of perturbation magnitude is shown in Fig. 9. It takes approximately 12 h for the spread to increase by a factor of 1.1, consistent with the spread growth for the 12-h data assimilation cycle shown in Fig 8. Between 24 and 48 h, latitudinally averaged spread grew from ~1.3 to ~2.0, with even greater growth between 48 and 72 h. This is consistent with the concept of a randomly oriented, small perturbation projecting more and more on the leading Lyapunov vectors as the control and perturbed are integrated forward (Toth and Kalnay 1993, Vannitsem and Nicolis 1997, Snyder and Hamill 2003).

Perhaps a deficiency of the additive noise perturbations that were used in the imperfect-model experiments was that they were not dynamically conditioned, i.e., they had no relevance to the flow of the day. Palmer (2002) has previously argued that the component of model error that is most important is the component that projects onto the growing forecast structures. Following a similar rationale, perhaps additive noise that were both consistent with model error statistics and dynamically conditioned to project to a greater extent onto the leading Lyapunov vectors of the system would produce analyses

18

with less error and greater spread growth. To test this, another additive noise experiment was performed with the imperfect model. In this experiment, instead of adding the noise samples directly at the time of the update, a slightly modified process was followed. First, the ensemble-mean analysis from 24 h prior was extracted. Additive noise perturbations were applied to the ensemble mean analysis, and 24-h forecasts were conducted.   The ensemble-mean forecast was subtracted to yield a set of evolved perturbations. After a global rescaling so their magnitude was consistent with that of the original additive noise perturbations, these evolved perturbations were used as the additive noise in the data assimilation.

Figure 10 shows the error, spread, and spread growth using these evolved perturbations. The minimum-error analysis is now ~1.56, compared to the ~1.62 previously, an ~3.7% decrease in error. The new minimum error now occurs at a larger localization radius, 4000 km, and a larger additive noise amount, 15-20%. Most notably, as Fig. 10c shows, the spread growth in the subsequent forecast has increased. For the parameters with the minimum error, spread growth is ~1.19 / 12-h cycle, which is actually larger than the spread growth for the perfect-model experiments at the same localization radius.   This suggests that for this model and experimental design, the structure of the additive errors and not model error were the primary cause of a deficiency of spread growth.

Figures 11 and 12 paint a more complex picture of the utility of evolved additive noise, however. Figure 11 presents the zonal-average RMS error and spread corresponding to the same 3000-km localization and 10 % additive noise used in Fig. 8.

19

While wind errors decreased minutely, tropical temperature errors substantially. The process of evolving the perturbations but then rescaling them so that they had the same global average perturbation size as the original additive noise enlarged the perturbations in the extratropics but shrunk them excessively in the tropics. Consequently, the unduly small temperature-error variances in the tropics prohibited drawing enough to the observations during the data assimilation, increasing the analysis error.

However, if we consider the analyses with 4000 km-localization and 15 % inflation (Fig. 12), the parameters where the evolved error was approximately at a minimum, there was a slight decrease in both extratropical winds and tropical temperature errors relative to the standard additive-noise simulations. However, there was an excess of spread, especially in the mid-latitudes. These results illustrate the difficulty of generating samples of the system noise in the data assimilation that are consistent with model errors statistics yet simultaneously project on growing forecast structures.

Did the evolved additive noise have a positive impact on longer-lead forecast error and spread? Figure 13 shows that the answer is tentatively yes. Three sets of 50-member ensemble spread and error curves are shown, averaged over 11 case days. The first set is for the data assimilation experiments shown in Figs. 8, with a 3000-km localization and 10% additive noise; call this "Add3000-10." The second uses the evolved perturbations but still uses the (sub-optimal) 3000-km localization and 10% additive noise as used in Fig. 11; call this "Evo3000-10." Finally, ensembles are produced using the lower-error 4000-km localization and 15% additive noise used in Fig.

12, or "Evo4000-15."   To test the statistical significance of changes in error, a 1000-sample paired block bootstrap between Add3000-10 and Evo4000-15 was performed for each forecast lead using their 11 daily global RMSE statistics, following Hamill (1999). The 5th and 95th percentiles of the resampled distribution are plotted atop the Evo4000-15 RMSE. There is a marginal positive impact on spread growth and error for Evo3000-10 relative to Add3000-10, but the Evo4000-15 retains the extra initial spread and grows the spread more quickly than Add3000-10, producing a greater spread-error consistency over the latter period of the forecast. Additionally, at long leads the forecast error is reduced, perhaps because the larger spread results in a more effective averaging of the ensemble. This difference for Evo4000-15 is statistically significant after day 4.5.

5. **Experiments with a T62 global forecast model**.

To test the results from the 2-layer primitive equation model in a more realistic setting, the National Centers for Environmental Prediction (NCEP) Global Forecast System (GFS) model was used with an EnSRF. Further details on the model and the data assimilation methodology were provided in Whitaker et al. (2008), with the following recent changes to the algorithm. An adaptive radiance bias correction algorithm developed by T. Miyoshi (personal communication) was included, which allows satellite radiances to be assimilated. The algorithm mimics what is done in the NCEP Grid-point statistical interpolation (GSI) variational system (Wu et al. 2002) and uses the same air-mass predictors in the bias calculation. Additionally, the fast parallel algorithm of Anderson and Collins (2007) was used to calculate the EnSRF increment.

The EnSRF assimilations were started on 0000 UTC 1 December 2007 and ended on 0000 UTC 10 January 2008. As in Whitaker et al. (2008), the NCEP GSI system was used for the forward operator calculation, and all conventional, satellite wind, and global positioning system radio-occultation data are assimilated, as well as Advanced Microwave Sounding Unit (AMSU) and High-Resolution Infrared Radiation Sounder (HIRS) radiances and Solar Backscatter Ultra-Violet Instrument (SBUV) ozone retrievals. Covariance localization with a length scale of 1500 km in the horizontal and 1.1 scale heights (in units of –ln(pressure)) was employed. Updates occurred every 6 h.

The EnSRF was run in two parallel cycles, the first employing scaled additive noise generated with 48-h minus 24-h forecast differences (the "NMC method;" Parrish and Derber 1992), and the second using the same scaled additive noise pre-evolved over the prior 24-h period. The scaling in both was 0.5. 10-member ensembles were conducted from each once daily from 0000 UTC initial conditions and integrated to 5 days lead. Figure 14 provides the results. The evolved additive noise starts with slightly higher spread, and that spread grows much faster during the first 24 hours of the forecast, so that at all subsequent leads the spread is significantly larger with the evolve additive noise. The ensemble-mean error is decreased slightly at the longest leads, but by an amount smaller than what was demonstrated with the 2-level model. Note, however, that the evolved ensemble was not "tuned." It is possible that a smaller error could have been achieved with, say, a longer localization radius for the evolved noise, as was the case for the 2-level model results. There is a notable inconsistency between spread and error at the longest leads, due presumably to the low model resolution, the strong model diffusion at this resolution, and the lack of any treatment of model error in this ensemble prediction

system. Nonetheless, the T62 GFS results suggest that the evolved additive noise will have a beneficial impact on spread growth during the early hours of the forecast, and it may provide some decrease in ensemble-mean error, especially at the longest leads.

6. **Conclusions**

While the EnKF has been demonstrated to be an advanced data assimilation method that can produce reduced-error initial conditions that are competitive with variational methods (Buehner et al. 2009ab), to date little experimentation has been performed on the characteristics of forecasts. The only center that currently runs the EnKF operationally, Environment Canada, has previously been concerned with the relatively slow growth of spread from their ensemble of initial conditions. This study attempted to determine whether the covariance localization, additive noise, or model error played the lead role in limiting spread growth from ensemble Kalman filters. The model chosen for these experiments was a T31, 2-level dry primitive equation global model. A uniform network of wind and temperature observations were assimilated using an ensemble square-root filter (EnSRF). This model and observation set was drastically simpler than are used in operational weather prediction. Still, this simplicity permitted a wide range of experiments to be conducted, and the model had some of the essential characteristics of more complex models, such as the ability to support internal gravity wave activity as well as baroclinic modes.

In perfect-model experiments, covariance localization was found to have a relatively modest effect on the growth of spread. In experiments where the EnSRF was also stabilized with covariance inflation, the localization reduced the growth of spread in

a global energy norm from approximately 1.204 / 12-h cycle with long localization scales to 1.175 when using a very short length scales. In comparison, at the longest localization length scale, changing from stabilizing the filter with covariance inflation to stabilizing it with additive noise reduced the spread growth from 1.204 to 1.16. The use of additive noise also increased the ensemble-mean analysis error substantially in the perfect-model experiment, with an error of ~0.17 m/s vs. ~0.12 for covariance inflation.

Next, a set of imperfect-model experiments was conducted using a T42 nature run with a different pole-to-equator temperature gradient and different mechanical damping time scale. Covariance inflation proved less useful for stabilizing perfect-model simulations, as previously discussed in Hamill and Whitaker (2005). Spread growth was also much smaller than in perfect-model experiments under stabilization by covariance inflation. Additive noise successfully stabilized the model, but spread growth was also much smaller, approximately 1.11 / 12-h cycle.

An examination of the characteristics of additive-noise perturbations showed that they typically grew very slowly during the first few hours of the forecast, but thereafter much more rapidly.   This suggested a possible improved ad-hoc procedure: instead of adding random additive noise samples, back up some period of time (in our study, 36 h), add the noise to an earlier ensemble mean, evolve the forecasts forward in time to the current update time, and use the rescaled, re-centered perturbations as dynamically conditioned additive noise. When this was done, this resulted in a modest (3.7%) decrease in analysis error, reduced the forecast error, and resulted in an improved spread-error consistency at the longest leads. L. Magnusson and M. Leutbecher (personal

communication, 2009) are also exploring the use of evolved additive noise for initializing ensemble predictions.

Results with a T62 version of the NCEP GFS provided confirmatory evidence that evolved additive noise could improve the rate of spread growth in the early hours of the ensemble forecasts, and possibly provide some reduction in ensemble-mean error, especially at the longest leads.

The application of evolved additive noise may appear at first glance somewhat impractical for operational numerical weather prediction, for costs of the EnSRF go up significantly, as evolving the additive noise increases the effective number of ensemble members that must be integrated forward in time during each data assimilation cycle. In higher-resolution, operational models, evolving the ensemble forward in time is the predominant computational expense. However, perhaps the additive noise could be evolved with a lower-resolution version of the forecast model, reducing its computational expense.

We end with some possible theoretical justification for flow-dependent additive noise samples. We hypothesize that flow-dependent noise may provide marginally more appropriate samples of the system error for two reasons. First, assume that system error is introduced at a constant rate during the x hours between EnKF updates. The system error introduced during the first hours will have a component that will project onto the system's leading Lyapunov vectors, and that part will grow like any other perturbation and thus be better represented by short-term evolved additive noise. Second, perhaps the actual system error was related to, say, an inappropriate estimate of mountain drag

(Shutts 2005). The standard additive noise perturbation may introduce noise over topography, regardless of whether there was strong flow in the region. Evolved additive noise will at least be more likely to decrease the amplitude of perturbations when the flow is weak and increase it when the flow is stronger. To the extent that model error is larger when and where the dynamics are more active, evolved additive noise should provide some improvement.

References

Aksoy, A., D. C. Dowell, C. Snyder, 2009: A multi-case comparative assessment of the ensemble Kalman filter for assimilation of radar observations. Part I: Storm-scale analyses. *Mon. Wea. Rev.*, **137**, 1805-1824.

Anderson, J. L., and S. L. Anderson, 1999: A Monte-Carlo implementation of the nonlinear filtering problem to produce ensemble assimilations and forecasts. *Mon. Wea. Rev.*, **127**, 2741-2758.

-------------, 2001: An ensemble adjustment filter for data assimilation. *Mon. Wea. Rev.*, **129**, 2884-2903.

-------------, and Collins, N., 2007: Scalable implementations of ensemble filter algorithms for data assimilation. *J. Atmos. Ocean. Tech.*, **24**, 1452-1463.

Bishop, C. H., and D. Hodyss, 2009a: Ensemble covariances adaptively localized with ECO-RAP. Part 1: tests on simple error models. *Tellus*, **61A**, 84-96.

-------------, and ---------------, 2009b: Ensemble covariances adaptively localized with ECO-RAP. Part 2: a strategy for the atmosphere. *Tellus*, **61A**, 97-111.

Buehner, M., and M. Charron, 2007: Spectral and spatial localization of background-error correlations for data assimilation. *Quart. J. Royal. Meteor. Soc.*, **133**, 615-630.

Buehner, M., P. L. Houtekamer, C. Charette, H. L. Mitchell, and B. He, 2009a: Intercomparison of Variational Data Assimilation and the Ensemble Kalman Filter for Global Deterministic NWP. Part I: Description of Configurations and Results from Idealized Experiments. *Mon. Wea. Rev.*, accepted pending minor revision. Available from Mark.Buehner@ec.gc.ca

27

------------, --------------, --------------, -------------, and --------------, 2009b:

> Intercomparison of Variational Data Assimilation and the Ensemble Kalman
> Filter for Global Deterministic NWP. Part II: Results from 1-Month Experiments
> with Real Observations. *Mon. Wea. Rev.*, accepted pending minor revision.
> Available from Mark.Buehner@ec.gc.ca.

Charron, M., G. Pellerin, L. Spacek, P. L. Houtekamer, N. Gagnon, H. L. Mitchell, and
> L. Michelin, 2009: Towards random sampling of model error in the Canadian
> ensemble prediction system. *Mon. Wea. Rev.*, in review. Available from
> Martin.Charron@ec.gc.ca.

Compo,G.P., J.S. Whitaker,and P.D. Sardeshmukh, 2006: Feasibility of a 100 year
> reanalysis using only surface pressure data. *Bull. Amer. Met. Soc.*, **87**, 175-190.

Courtier, P., J.-N. Thepaut, and A. Hollingsworth, 1994: A strategy for operational
> implementation of 4D-Var, using an incremental approach. *Quart. J. Royal*
> *Meteor. Soc.*, **120**, 1367-1387.

Ehrendorfer, M., 2007: A review of issues in ensemble Kalman filtering. *Meteor.*
> *Zeitschrift*, **16**, 795-818.

Evensen, G., 1994: Sequential data assimilation with an nonlinear quasi-geostrophic
> model using Monte Carlo methods to forecast error statistics. *J. Geophys. Res.*,
> **99(C5)**, 10143-10162.

Evensen, G., 2006: *Data Assimilation, the Ensemble Kalman Filter*. Springer Press, 285
> pp.

Gaspari, G., and S. E. Cohn, 1999: Construction of correlation functions in two and three
> dimensions. *Quart. J. Royal Meteor. Soc.*, **125**, 723-757.

Hamill, T. M., 1999: Hypothesis tests for evaluating numerical precipitation forecasts. *Wea. Forecasting*, **14**, 155-167.

--------------, and C. Snyder, 2000: A hybrid ensemble Kalman filter / 3d-variational analysis scheme. *Mon. Wea. Rev.*, **128**, 2905-2919.

--------------, J. S. Whitaker, and C. Snyder, 2001: Distance-dependent filtering of background error covariance estimate in an ensemble Kalman filter. *Mon. Wea. Rev.*, **129**, 2776-2790.

--------------, C. Snyder, and R. Morss, 2002: Analysis-error statistics of a quasigeostrophic model using three-dimensional variational assimilation. *Mon. Wea. Rev.*, **130**, 2777-2790.

--------------, and J. S. Whitaker, 2005: Accounting for the error due to unresolved scales in ensemble data assimilation: a comparison of different approaches. *Mon. Wea. Rev.*, **133**, 3132-3147.

--------------, 2006: Ensemble-based atmospheric data assimilation. Chapter 6 in *Predictability of Weather and Climate*, T. N. Palmer and R. Hagedorn, Eds. Cambridge Press, 124-156.

Houtekamer, P. L., and H. L. Mitchell, 1998: Data assimilation using an ensemble Kalman filter technique. *Mon. Wea. Rev.*, **126**, 796-811.

--------------, ----------------, 2001: A sequential ensemble Kalman filter for atmospheric data assimilation. *Mon. Wea. Rev.*, **129**, 123-137.

--------------, ----------------, G. Pellerin, M. Buehner, M. Charron, L. Spacek, and B. Hansen, 2005: Atmospheric data assimilation with an ensemble Kalman filter: results with real observations. *Mon. Wea. Rev.*, **133**, 604-620.

--------------, ----------------, 2005: Ensemble Kalman filtering. *Quart. J. Roy. Meteor. Soc.*, **131**, 3269–3289.

--------------, ----------------, and X. Deng, 2009: Model error representation in an operational ensemble Kalman filter. *Mon. Wea. Rev.*, **137**, 2126-2143.

Hunt, B. R., E. J. Kostelich, and I. Szunyogh, 2006: Efficient data assimilation for spatiotemporal chaos: A local ensemble transform Kalman filter. *Physica D*, **230**, 112-126. http://dx.doi.org/10.1016/j.physd.2006.11.008

Kalnay, E., B. Hunt, E. Ott, and I. Szunyogh, 2006: Ensemble forecasting and data assimilation: two problems with the same solution? Chapter 7 of *Predictability of Weather and Climate*, T. N. Palmer and R. Hagedorn, Eds., Cambridge Press, 157-180.

Kepert, J. D., 2009: Covariance localization and balance in an ensemble Kalman filter. *Quart. J. Royal. Meteor. Soc*., **135**, 1157-1176.

Le-Dimet, F.-X., and O. Talagrand, 1986: Variational algorithms for analysis and assimilation of meteorological observations. Theoretical aspects. *Tellus*, **38A**, 97-110.

Lee, S., and I. M. Held, 1993: Baroclinic wave packets in models and observations. *J. Atmos. Sci.*, **50**, 1413-1428.

Lorenc, A. C., 2003: The potential of the ensemble Kalman filter for NWP – a comparison with 4D-Var. *Quart. J. Royal Meteor. Soc*., **129**, 3183-3203.

Lorenz, E. N., 1993: *The Essence of Chao*s. University of Washington Press, 227 pp.

Meng, Z., and F. Zhang, 2008a: Test of an ensemble Kalman filter for mesoscale and regional-scale data assimilation. Part III: Comparison with 3DVar in a real-data case study. *Mon. Wea. Rev.*, **136**, 522-540.

--------------, and --------------, 2008b: Test of an ensemble Kalman filter for mesoscale and regional-scale data assimilation. Part IV: Comparison with 3DVar in a month-long experiment. *Mon. Wea. Rev.*, **136**, 3671-3682.

Mitchell, H. L., and P. L. Houtekamer, 2000: An adaptive ensemble Kalman filter. *Mon. Wea. Rev.,* **128**, 416-433.

Mitchell, H. L., P. L. Houtekamer, and G. Pellerin, 2002: Ensemble size, balance, and model-error representation in an ensemble Kalman filter. *Mon. Wea. Rev*., **130**, 2791-2808.

Miyoshi, T., and S. Yamane, 2007: Local Ensemble Transform Kalman Filtering with an AGCM at a T159/L48 Resolution. *Mon. Wea. Rev.*, **135**, 3841–3861.

Palmer, T. N., 2002: Predicting uncertainty in numerical forecasts. Chapter 1 of *Meteorology at the Millennium, Volume 83, International Geophysics*. Academic Press, 450 pp.

Parrish D., and J. Derber, 1992: The National Meteorological Center's spectral statistical-interpolation analysis system. *Mon. Wea. Rev.*, **120**, 1747–1763.

Rabier, F., H. Jarvinen, E. Klinker, J.-F. Mahfouf, and A. Simmons, 2000: The ECMWF operational implementation of four-dimensional variational assimilation: I: experimental results with simplified physics. *Quart. J. Royal Meteor. Soc*., **126**, 1143-1170.

Shutts, G. J., 2005: A kinetic energy backscatter algorithm for use in ensemble

      prediction systems. *Quart. J. Royal Meteor. Soc*., **132**, 3079-3102.

Snyder, C., and T. M. Hamill, 2003: Leading Lyapunov vectors of a turbulent baroclinic

      jet in a quasigeostrophic model. *J. Atmos. Sci*., **60**, 683-688.

Szunyogh, I., E. J. Kostelich, G. Gyarmati, E. Kalnay, B. R. Hunt, E. Ott, E. Satterfield,

      and J. A. Yorke, 2008: A local ensemble transform Kalman filter data

      assimilation system for the NCEP global model. *Tellus,* **60A,** 113–130.

Toth, Z., and E. Kalnay, 1993: Ensemble forecasting at NMC: the generation of

      perturbations. *Bull. Amer. Meteor. Soc*., **74**, 2317-2330.

Torn, R. D., and G. J. Hakim, 2008: Performance characteristics of a pseudo-operational

      ensemble Kalman filter. *Mon. Wea.Rev.,* **136,** 3947–3963.

--------------, and --------------, 2009: Ensemble Data Assimilation applied to RAINEX

      observations of Hurricane Katrina (2005). *Mon. Wea. Rev*., **137**, 2817-2829.

Vannitsem, S. and C. Nicolis, 1997: Lyapunov vectors and error growth patterns in a

      T21L3 quasigeostrophic model. *J. Atmos. Sci*., **54**, 347-361.

Wang, X., D. Barker, C. Snyder, T. M. Hamill, 2008: A hybrid ETKF-3DVAR data

      assimilation scheme for the WRF model. Part II: real observation experiments.

      Mon. Wea. Rev., **13**6, 5116-5131.

Whitaker, J. S., and T. M. Hamill, 2002: Ensemble data assimilation without perturbed

      observations. *Mon. Wea. Rev*., **130**, 1913-1924.

---------------, G.P. Compo, X. Wei and and T.M. Hamill, 2004: Reanalysis without

      radiosondes using ensemble data assimilation. *Mon. Wea. Rev.*, **132**, 1190-1200.

---------------, T. M. Hamill, X. Wei, Y. Song, and Z. Toth, 2008: Ensemble data

assimilation with the NCEP global forecast system. *Mon. Wea. R*ev., **136**, 463–482.

----------------, G. P. Compo, and J.-N. Thépaut, 2009: A comparison of variational and ensemble-based data assimilation systems for reanalysis of sparse observations. *Mon. Wea. Rev.*, **137**, 1991-1999.

Wu, W.S., R.J. Purser, and D.F. Parrish, 2002: Three-dimensional variational analysis with spatially inhomogeneous covariances. *Mon. Wea. Rev.*, **130**, 2905–2916.

Zhang, F., C. Snyder, and R. Rotunno, 2003: Effects of moist convection on mesoscale predictability. *J. Atmos. Sci.*, **60**, 1173-1185.

----------------, Y. Weng, Z. Meng, J. A. Sippel and C. H. Bishop, 2009: Cloud-resolving hurricane initialization and prediction through assimilation of Doppler radar observations with an ensemble Kalman Filter: Humberto (2007). *Mon. Wea. Rev.*, **137**, 2105-2125.

LIST OF FIGURES

ensemble spread growth rate over the 12-h period between data assimilation updates.

Figure 6. (a) Zonal and time average of the upper-level u-wind component analysis error and spread for the 3000 km localization / 50% inflation rate data point shown in Fig. 5. (b) As in (a), but for lower-level u-wind component. (c) As in (a), but for interface potential temperature.

Figure 7: As in Fig. 5, but for imperfect-model experiments stabilized by additive noise. Dot indicates the localization radius and additive noise amount where error was approximately at a minimum and spread was consistent with error. Spread and error for this combination is examined more in Fig. 8.

Figure 8: As in Fig. 6, but for the 3000 km localization and 10% additive noise experiment data shown in Fig. 7.

Figure 9: Growth of small model-error additive noise perturbations around a state sampled from the T31 forecast model nature run. Dashed line indicates the relative proportion of total energy in the perturbations at the initial time. Perturbations are scaled so that their average magnitude at each latitude is 1.0 at the initial time.

Figure 10. As in Fig. 7, but where evolved additive noise is used. Black and grey dots indicate the localization radius and additive noise used for plots in Figs 11 and 12, respectively.
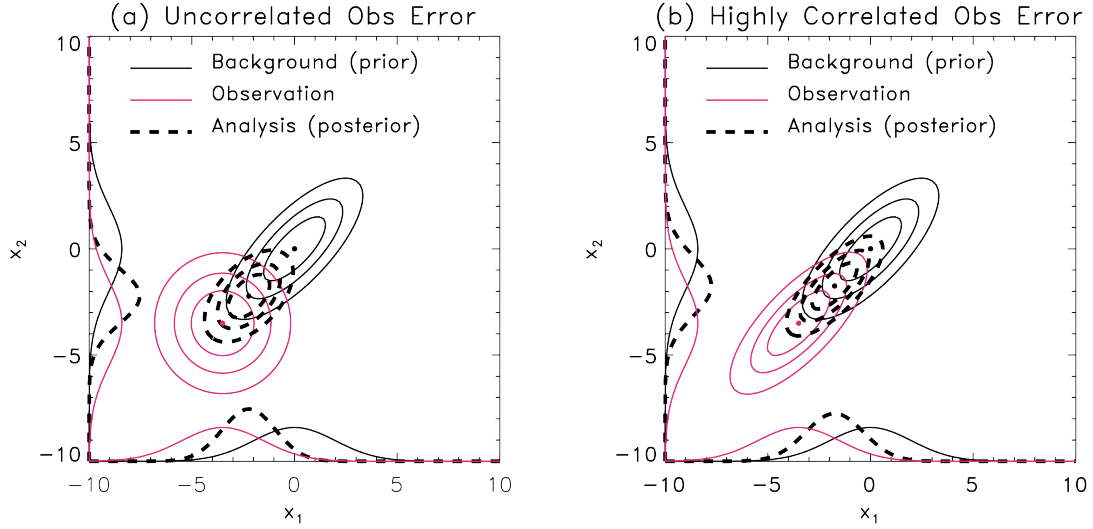
Figure 11. As in Fig. 8, but where evolved additive noise is used. Data is plotted for the localization of 3000 km and evolved additive noise magnitude of 10%, the black

dot in Fig. 10. Solid grey line repeats the RMS error line from Fig. 8 for comparison.

Figure 12. As in Fig. 8, but where evolved additive noise is used. Data is plotted for the localization of 4000 km and evolved additive noise magnitude of 20%, the grey dot in Fig. 10. Solid grey line repeats the RMS error line from Fig. 8 for comparison.

Figure 13: Ensemble forecast spread and RMS error in the energy norm, initialized from imperfect-model additive-noise forecasts with 3000 km localization and 10% additive noise scaling (Add3000-10, black lines); from evolved additive noise with 3000 km localization and 10% additive noise scaling (Evo3000-10, red lines); and evolved additive noise with 4000 km localization and 15% additive noise scaling (Evo4000-15, red lines). Error bars represent the 5th and 95th percentiles from a paired block bootstrap between the Add3000-10 and the Evo4000-15.

Figure 14: Ensemble forecast spread of mean-sea level pressure (dashed lines) and error (solid lines) from T62 GFS experiments with evolved additive noise (red lines) and conventional additive noise (black lines). 5th and 95th percentiles of a block bootstrap assuming independence of samples on each day are over-plotted on the evolved additive noise lines.

Figure 1: Illustration of how an improper assumption of independent observations may result in a posterior with an underestimate of covariances. Prior covariance and marginal distributions are shown in solid black lines; the three contours encompass 25, 50, and 75 percent of the probability density. Observation-error covariances and marginals are shown in solid red, and posterior distribution in dashed black after the application of conventional data assimilation techniques. Panel (a) shows the distributions when the observations are assumed to have uncorrelated errors, panel (b) with correlated errors.
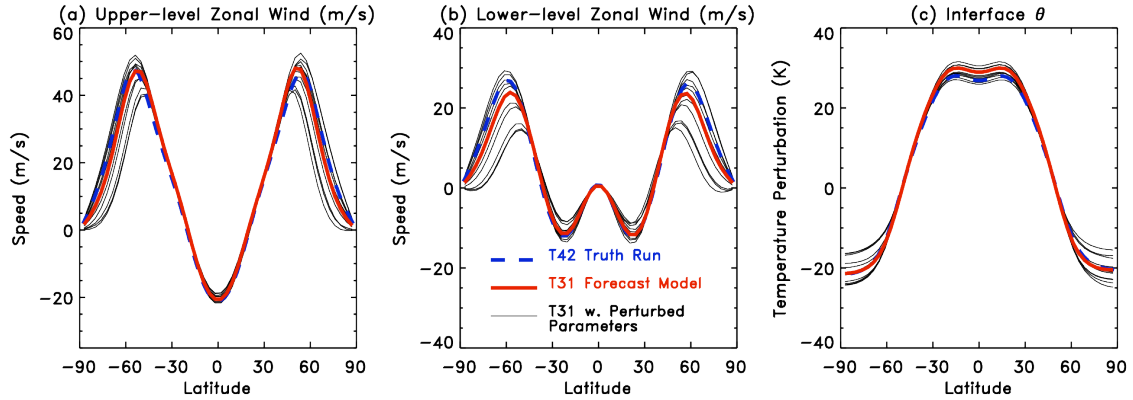
Figure 2: Zonal-mean averages for (a) upper-level u-wind component, (b) lower-level u-wind component, and (c) interface barotropic potential temperature in nature runs. Red line indicates the T31 model used for data assimilation experiments, dashed blue line the T42 model nature run used in imperfect-model experiments, and thin solid lines the various T31 models run with perturbed pole-to-equator temperature difference and perturbed mechanical damping time scales.
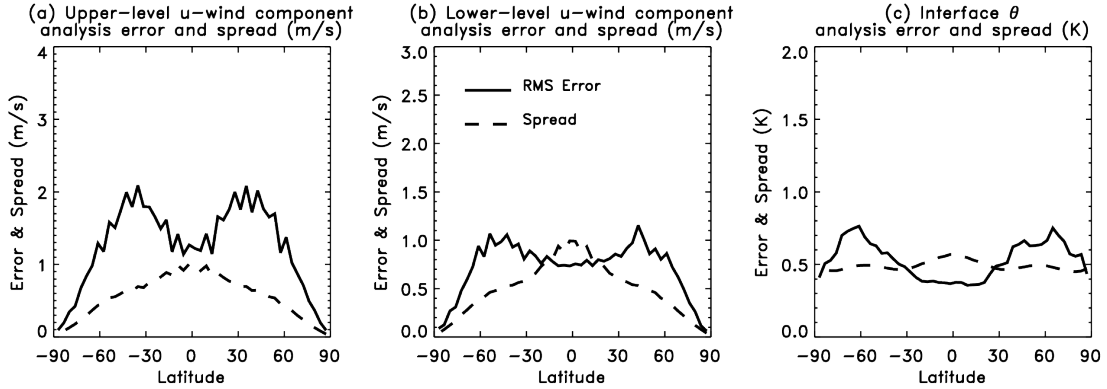
Perfect-model
Analysis Error & Spread, Energy Norm
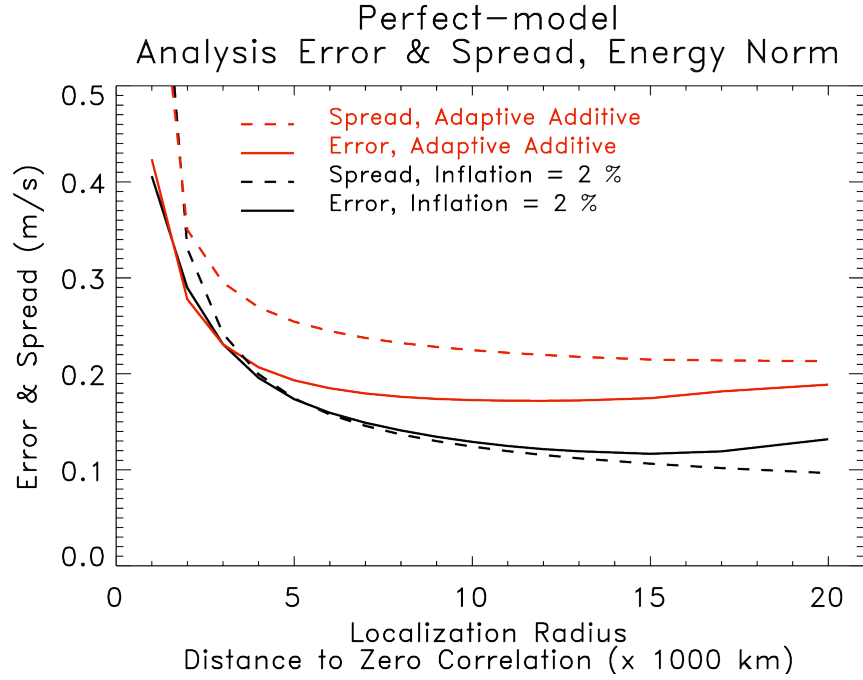
Figure 3: Spread and error for perfect-model experiments when ensemble is stabilized by covariance inflation (black curves) and adaptive additive noise (red curves).
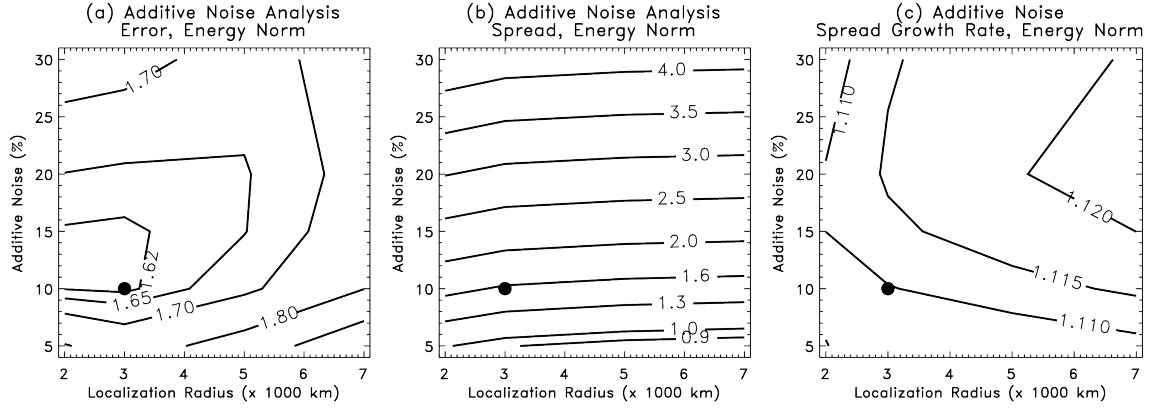

Perfect-Model, Growth of
Ensemble Spread, Energy Norm
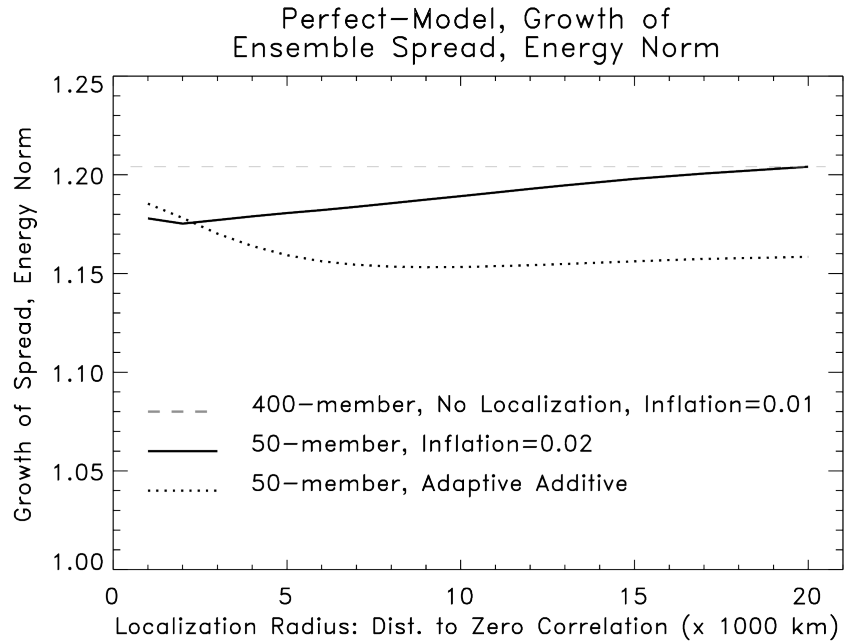
Figure 4: Growth of ensemble forecast spread during the 12-h between update steps.
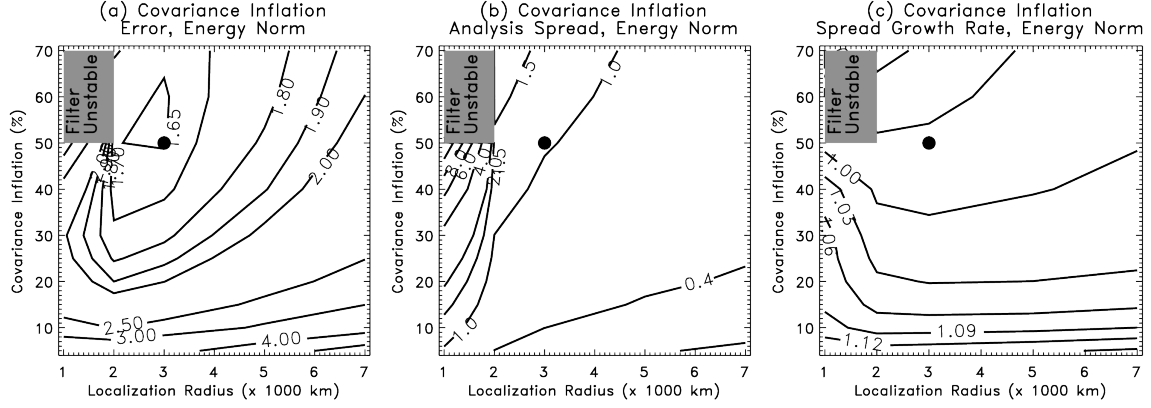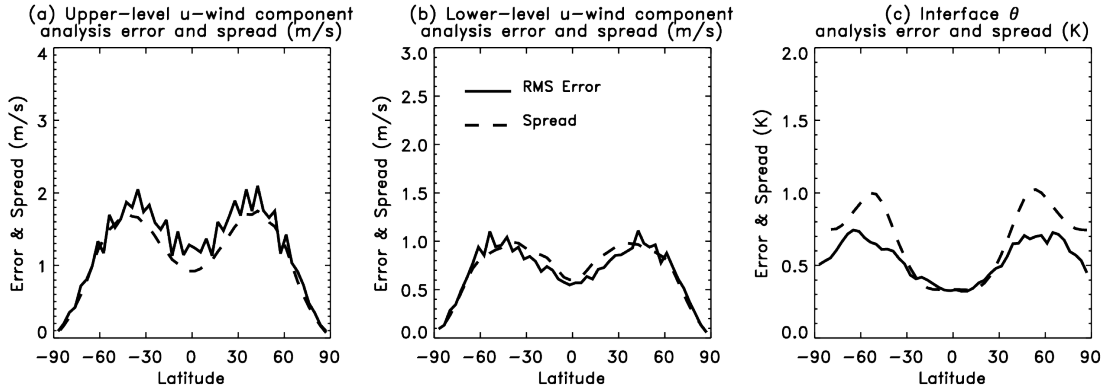
Figure 5. (a) Ensemble-mean RMS error in the energy norm for imperfect-model experiments with the EnSRF. Data is plotted as a function of the covariance localization radius (x axis) and the amount of covariance inflation (y axis). The black dot indicates the (localization radius, inflation amount) pair used in Fig. 6. (b) As in (a), but for ensemble spread in the energy norm. (c) As in (a), but for the ensemble spread growth rate over the 12-h period between data assimilation updates.

Figure 6. (a) Zonal and time average of the upper-level *u*-wind component analysis error and spread for the 3000 km localization / 50% inflation rate data point shown in Fig. 5. (b) As in (a), but for lower-level *u*-wind component. (c) As in (a), but for interface potential temperature.
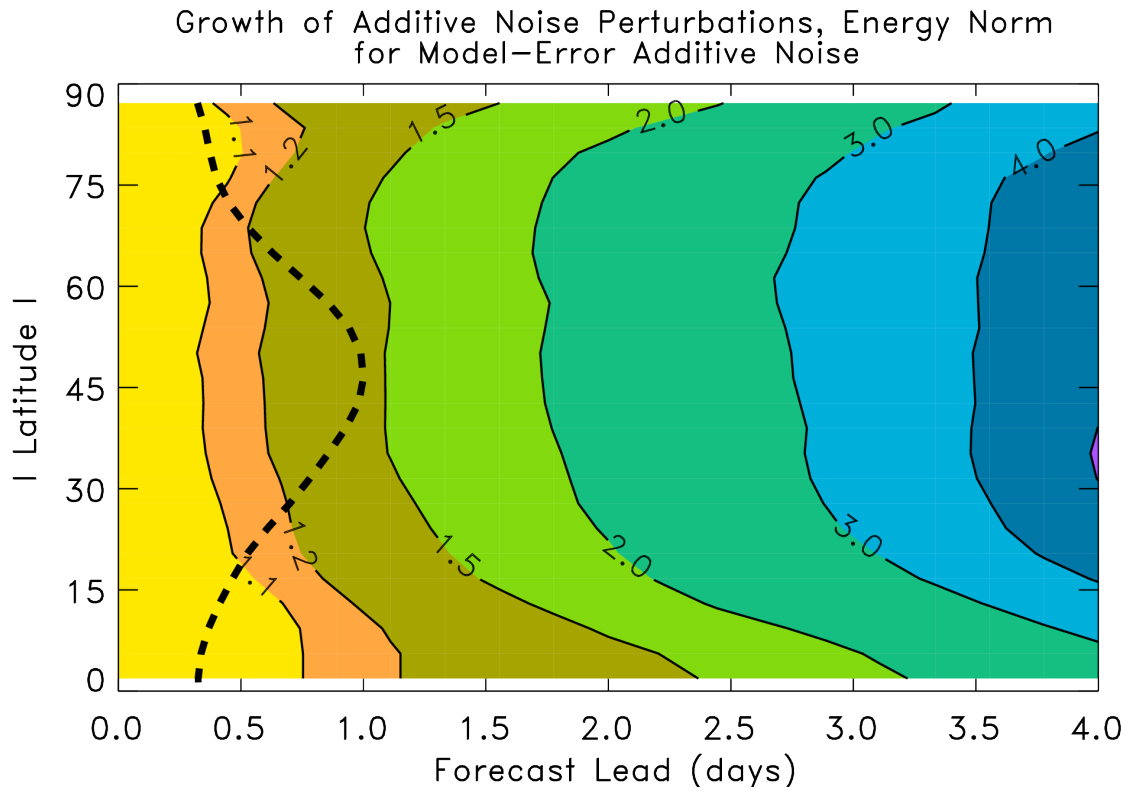
Figure 7: As in Fig. 5, but for imperfect-model experiments stabilized by additive noise. Dot indicates the localization radius and additive noise amount where error was approximately at a minimum and spread was consistent with error. Spread and error for this combination is examined more in Fig. 8.
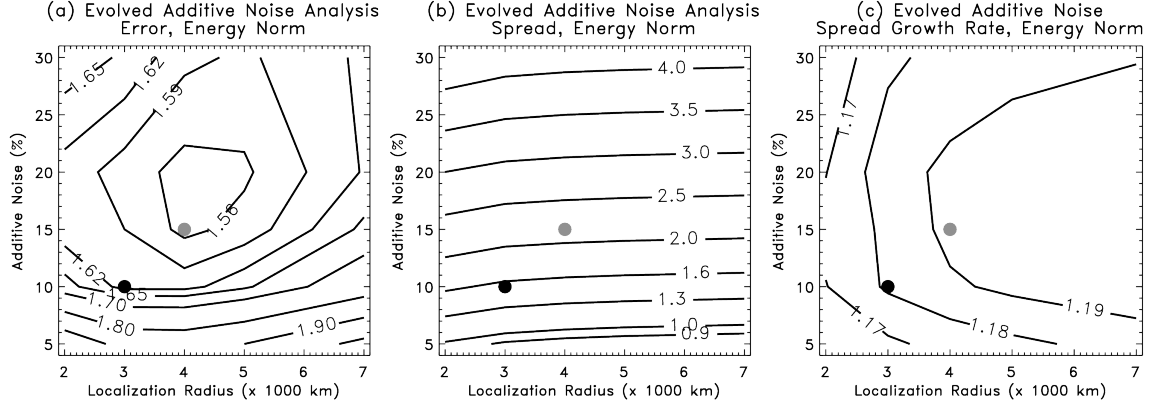

Figure 8: As in Fig. 6, but for the 3000 km localization and 10% additive noise experiment data shown in Fig. 7.
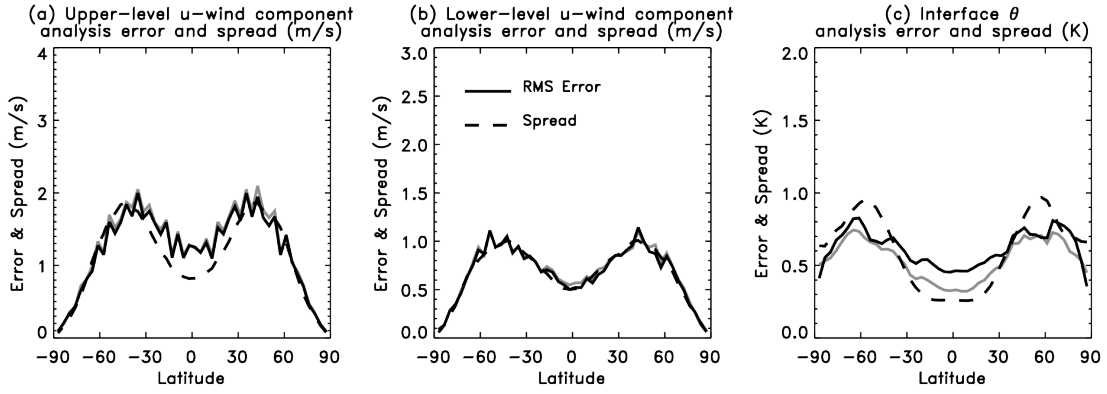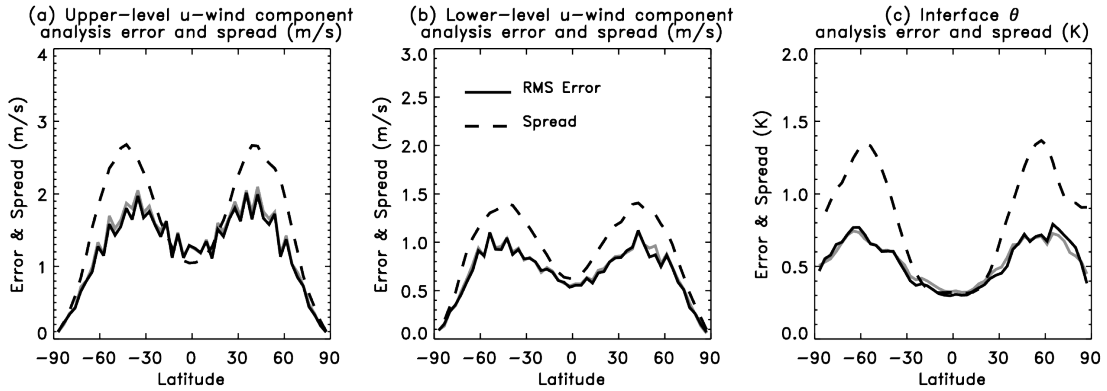
Figure 9: Growth of small model-error additive noise perturbations around a state sampled from the T31 forecast model nature run. Dashed line indicates the relative proportion of total energy in the perturbations at the initial time. Perturbations are scaled so that their average magnitude at each latitude is 1.0 at the initial time.
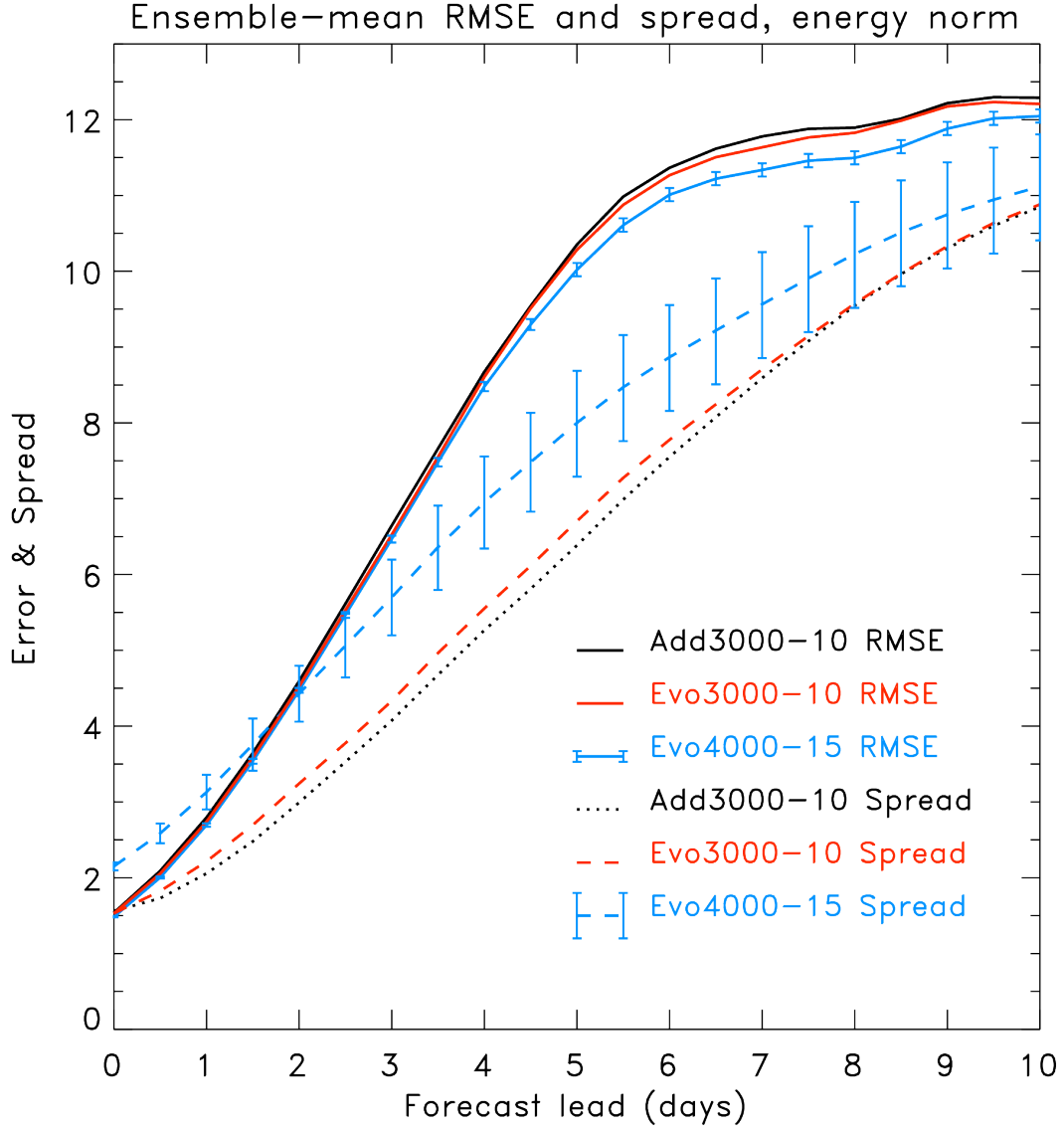
(a) Evolved Additive Noise Analysis Error, Energy Norm

(b) Evolved Additive Noise Analysis Spread, Energy Norm

(c) Evolved Additive Noise Spread Growth Rate, Energy Norm

Figure 10. As in Fig. 7, but where evolved additive noise is used. Black and grey dots indicate the localization radius and additive noise used for plots in Figs 11 and 12, respectively.
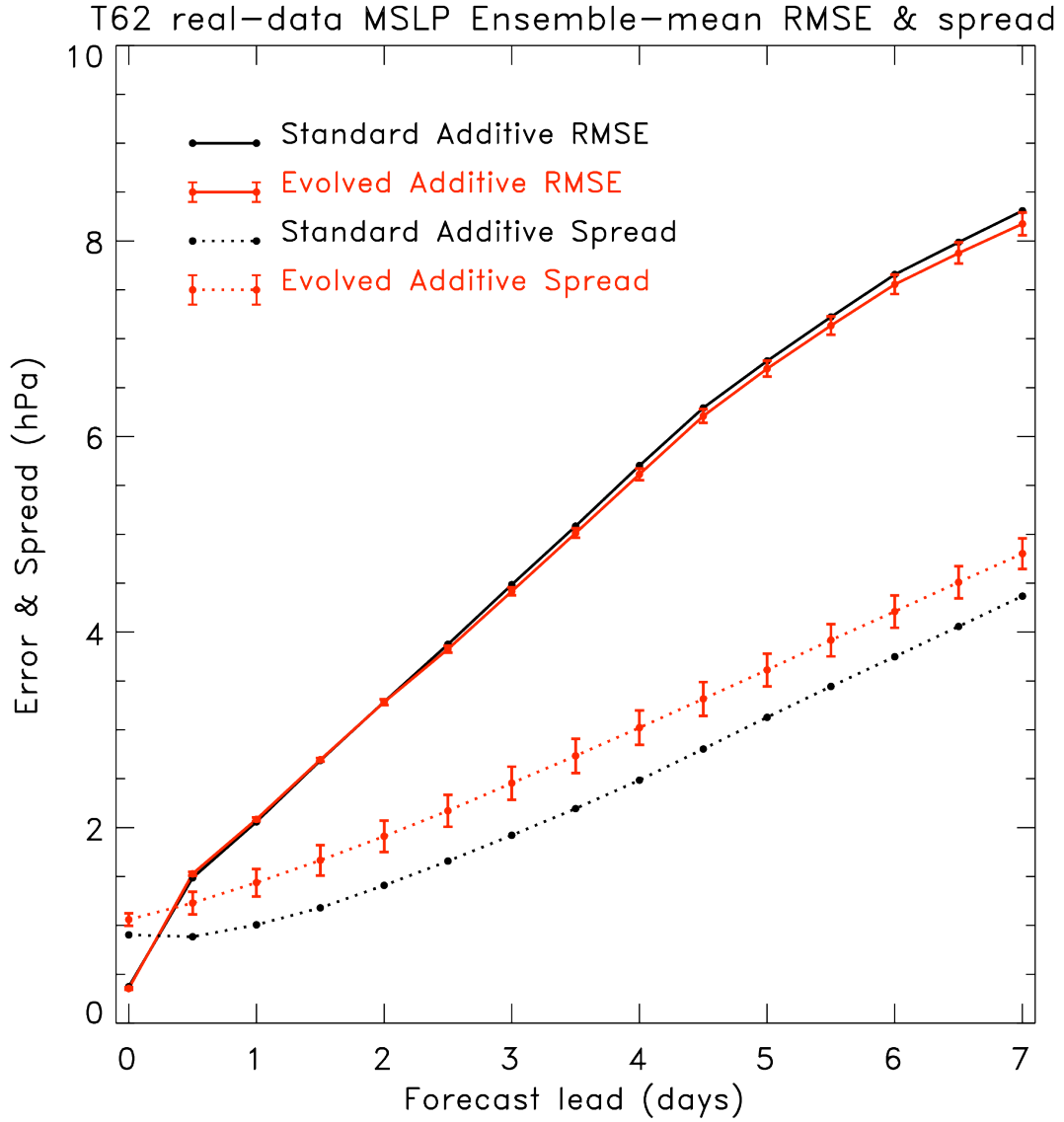

(a) Upper-level u-wind component analysis error and spread (m/s)

(b) Lower-level u-wind component analysis error and spread (m/s)

(c) Interface θ analysis error and spread (K)

Figure 11. As in Fig. 8, but where evolved additive noise is used. Data is plotted for the localization of 3000 km and evolved additive noise magnitude of 10%, the black dot in Fig. 10. Solid grey line repeats the RMS error line from Fig. 8 for comparison.

(a) Upper-level u-wind component analysis error and spread (m/s)

(b) Lower-level u-wind component analysis error and spread (m/s)

(c) Interface θ analysis error and spread (K)

Figure 12. As in Fig. 8, but where evolved additive noise is used. Data is plotted for the localization of 4000 km and evolved additive noise magnitude of 20%, the grey dot in Fig. 10. Solid grey line repeats the RMS error line from Fig. 8 for comparison.

42

Figure 13: Ensemble forecast spread and RMS error in the energy norm, initialized from imperfect-model additive-noise forecasts with 3000 km localization and 10% additive noise scaling (Add3000-10, black lines); from evolved additive noise with 3000 km localization and 10% additive noise scaling (Evo3000-10, red lines); and evolved additive noise with 4000 km localization and 15% additive noise scaling (Evo4000-15, red lines). Error bars represent the 5th and 95th percentiles from a paired block bootstrap between the Add3000-10 and the Evo4000-15.

Figure 14: Ensemble forecast spread of mean-sea level pressure (dashed lines) and error (solid lines) from T62 GFS experiments with evolved additive noise (red lines) and conventional additive noise (black lines). 5th and 95th percentiles of a block bootstrap assuming independence of samples on each day are over-plotted on the evolved additive noise lines.